

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-078640

(43)Date of publication of application : 14.03.2003

(51)Int.Cl.

H04M 11/00
G06K 9/00
G06K 9/03
H04B 7/26
H04M 1/00

(21)Application number : 2001-268194

(71)Applicant : HITACHI LTD

(22)Date of filing : 05.09.2001

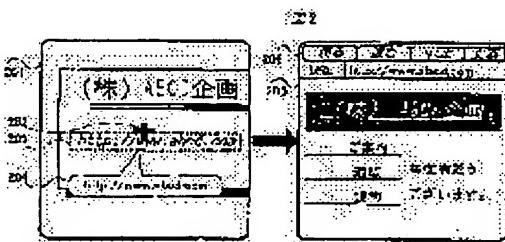
(72)Inventor : KAGEHIRO TATSUHIKO
SEKI MINENOB
SAKO YUTAKA

(54) PORTABLE TERMINAL EQUIPMENT AND COMMUNICATION SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a portable terminal for presenting information collected by using the character recognized result in the portable terminal, with which an imaging device is incorporated and a character recognizing function is added.

SOLUTION: The passage of character line extraction is displayed on the display part of the portable terminal as an occasion demands and resolution of an image to be inputted to recognizing processing is improved. Thus, an operator can easily select a target character line and by improving the resolution, the portable terminal improved in character recognition accuracy is provided.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-78640

(P2003-78640A)

(43) 公開日 平成15年3月14日 (2003.3.14)

(51) Int.Cl.⁷
H 04 M 11/00
G 06 K 9/00
9/03
H 04 B 7/26
H 04 M 1/00

識別記号
3 0 2

F I
H 04 M 11/00
G 06 K 9/00
9/03
H 04 M 1/00
H 04 B 7/26

テマコード(参考)
3 0 2 5 B 0 6 4
S 5 K 0 2 7
J 5 K 0 6 7
R 5 K 1 0 1
M

審査請求 未請求 請求項の数10 O L (全 9 頁)

(21) 出願番号 特願2001-268194(P2001-268194)

(22) 出願日 平成13年9月5日 (2001.9.5)

(71) 出願人 000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地

(72) 発明者 影広 達彦
東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(72) 発明者 関 峰伸
東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(74) 代理人 100075096
弁理士 作田 康夫

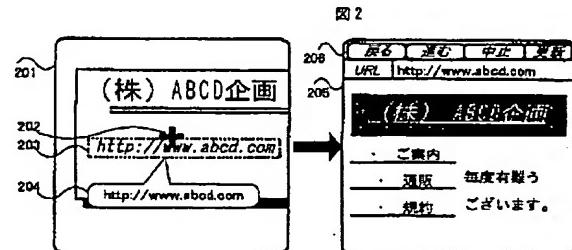
最終頁に続く

(54) 【発明の名称】 携帯端末装置及び通信システム

(57) 【要約】

【課題】撮像素子を内蔵し文字認識機能を付加した携帯端末であって、該文字認識結果を用いて収集した情報を提示する携帯端末を提供すること。

【解決手段】携帯端末の表示部に文字行抽出経過を隨時表示し、認識処理に入力する画像を高解像度化することで、操作者が容易に対象文字行を選択する事が可能であって、かつ、高解像度化する事により文字認識精度が向上した携帯端末を提供する。



【特許請求の範囲】

【請求項1】画像を取得する手段と、上記取得画像を表示する手段と、上記取得画像中の文字行記載領域を抽出する手段と、上記抽出手段によって抽出される文字行記載領域を確定する手段と、上記確定された文字行記載領域に記載される文字行を認識する手段と、上記認識結果をネットワークを介して転送する手段と、上記認識結果に基づくコンテンツを受信する手段とを有する携帯端末であって、上記抽出手段は、上記確定手段に入力があるまで上記文字行記載領域を変動させながら連続して抽出し、かつ上記抽出した文字行記載領域を上記表示手段上に常に表示することを特徴とする携帯端末装置。

【請求項2】上記表示手段はマーカーを有し、上記抽出手段は上記取得画像中のパターンを検出し、上記マーカーに最近傍の上記パターンを核として近接する上記パターンを連結することで文字行記載領域を抽出することを特徴とする請求項1記載の携帯端末装置。

【請求項3】上記マーカーは移動可能であることを特徴とする請求項2に記載の携帯端末装置。

【請求項4】上記携帯端末はさらに、上記文字行認識手段によって認識された複数の認識結果を提示する手段と、上記複数の認識結果のうちの1を選択する手段を有することを特徴とする請求項1乃至3のいずれかに記載の携帯端末装置。

【請求項5】上記携帯端末はさらに高精細化手段を有し、上記高精細化手段は、上記確定された文字行記載領域を画素補完し、上記画素補完された部分画像を平滑化し、さらに2値化することを特徴とする請求項1乃至4のいずれかに記載の携帯端末装置。

【請求項6】上記パターンは上記取得画像中の連結成分の外接矩形であることを特徴とする請求項2乃至5のいずれかに記載の携帯端末装置。

【請求項7】携帯端末と、ネットワークを介して上記携帯端末と接続される計算機からなる通信システムであつて、上記携帯端末は、画像取得手段と該取得画像の表示手段と、上記取得画像中の文字行記載領域抽出手段と、上記抽出手段によって抽出された文字行記載領域を確定する手段と、上記確定された領域中の文字行認識を行う手段と、上記認識結果を上記ネットワークを介して上記計算機に転送する手段とを有し、上記文字行認識手段は、上記確定手段に入力があるまで上記文字行記載領域を変動させながら継続抽出し、かつ上記抽出した文字行記載領域を上記表示手段上に常に表示し続け、上記計算機は上記認識結果をもとにネットワーク上の情報検索をおこなって、検索結果を上記携帯端末に転送する手段を有することを特徴とする通信システム。

【請求項8】携帯端末と、ネットワークを介して上記携帯端末と接続される計算機からなる通信システムであつて、上記携帯端末は、画像取得手段と該取得画像の表示手段と、上記取得画像中の文字行記載領域抽出手段と、

上記抽出手段によって抽出される領域を確定する手段と、上記確定手段によって確定された画像を上記ネットワークを介して上記計算機に転送する手段とを有し、上記文字行認識手段は、上記確定手段に入力があるまで上記文字行記載領域を変動させて継続抽出し、かつ上記抽出した文字行記載領域を上記表示手段上に常に表示し、上記計算機は、上記転送された画像の文字認識を行う手段と、上記認識結果をもとにネットワーク上の情報検索をおこなって、検索結果を上記携帯端末に転送する手段を有することを特徴とする通信システム。

【請求項9】上記計算機は、上記取得した認識結果がURLアドレス、メールアドレス、もしくは他の文字列であるかを判別する手段を有し、さらに、上記判別結果がURLアドレスであった場合には、該URLアドレスの示すコンテンツを端末に送信する手段と、上記判別結果がメールアドレスであった場合には、メール作成用のプログラムを上記端末に送信する手段と、上記判別結果がその他の文字列であった場合には、該文字列をもとに検索プログラムを起動して上記ネットワーク内の情報を検索し、該検索結果を上記端末に送信する手段とを有することを特徴とする請求項7又は8のいずれかに記載の通信システム。

【請求項10】上記計算機から転送される上記検索結果を上記携帯端末が受信する際に課金を行う手段を有することを特徴とする請求項7乃至9のいずれかに記載の通信システム。

【発明の詳細な説明】

【0001】

【発明が属する技術分野】今後、携帯端末や携帯電話に内蔵するCPUの高性能化や、CCD等の低価格化によって、撮像素子を内蔵した携帯端末が普及する。撮像素子を内蔵した携帯端末を用いて、ユーザーが着目した文字列を含んだ景観画像を採取し、その文字列を認識することにより、ネットワーク上から関連情報を引き出す事が可能になる。景観画像中から目的の文字列を選択するためには、携帯端末上の限られたユーザーインターフェースを用い、かつ、操作者が直感的に理解し易い指示方法が必要である。また、携帯端末には構造上、高解像度画像の採取可能なデバイスを格納することは難しい。そのため携帯端末のカメラで採取した低解像度の画像を用いて、文字認識を行う。しかし、解像度と認識精度はトレードオフの関係にあり、高い認識精度を要求する場合、入力画像を加工する必要がある。

【0002】

【従来の技術】景観画像中からの文字認識結果を用いた情報検索は、特開2000-331006、特開平10-254622に記載がある。また、特開2000-23012は携帯端末を用いた翻訳機について記載されており、この発明中のGUIに対象文字行を当てはめる線分が画像表示部に表示される。しかし、線分は固定で、操作者が景観画像を採取する際の目

安として利用するのみで、文字行の位置やサイズを操作者が合わせる必要がある。また、携帯端末内の低解像度撮像系で採取した画像を対象とする文字認識に関する発明は存在しない。

【0003】

【発明が解決しようとする課題】携帯電話または携帯端末にCCDカメラを内蔵し、景観画像をキャプチャして文字認識を行い、翻訳機能、検索機能などを搭載する手段は、特開2000-23012で述べられている。しかし、操作者が景観画像中から対象文字行を容易に選択する手段は存在しない。そこで本発明は、操作者に対し入力画像中からの文字行抽出結果を随時提示し、任意の対象文字行を容易に選択する事を可能にする。また、携帯端末などに内蔵する撮像素子は、コストや設置容積の点から、高解像度の画像を採取可能に設計することは困難である。低解像度の画像を用いて認識処理を行うと、認識精度に悪影響が生じ、操作者の目的が阻害される。そこで本発明では、採取した低解像度の画像から、文字行位置の抽出処理を行った後、その文字行位置の部分画像に対し拡大処理を行うことにより、文字行部分画像を高解像度化し、認識精度の向上を図る。

【0004】

【課題を解決するための手段】本願において開示される発明のうち、代表的なものの概要は以下の通りである。取得画像中の文字行記載領域を抽出する手段を有する携帯端末であって、確定手段に入力があるまで上記文字行記載領域を変動させながら連続して抽出し、かつ上記抽出した文字行記載領域を表示手段上に常に表示することで抽出する文字行記載領域を確定し、上記確定された文字行記載領域に記載される文字行を認識する手段と、上記認識結果をネットワークを介して転送する手段と、上記認識結果に基づくコンテンツを受信する手段とを有する携帯端末。さらに取得画像の高精細化手段を有し、上記確定された文字行記載領域を画素補完し、上記画素補完された部分画像を平滑化し、さらに2値化を行うことを特徴とする携帯端末。又、上記携帯端末と、ネットワークを介して上記携帯端末と接続される計算機からなる通信システムであって、上記形態端末で取得した認識結果若しくは上記文字行抽出領域を上記ネットワークを介して上記計算機に転送し、上記計算機が転送された上記認識結果若しくは上記文字行抽出領域の情報をもとにネットワーク上の情報検索をおこなって、検索結果を上記携帯端末に転送する手段を有することを特徴とする通信システム。

【0005】

【発明の実施の形態】図1に本発明に用いる携帯端末と、入力データとなる文字行が記載された物体を示している。101は、携帯端末の1形態であり、102に画像表示部があり、反対側の面には撮像素子を用いたカメラが格納されている。103は操作レバーで、104はボタンである。

る。また画像表示部はタッチパネルになっており、操作者が直接指示できるようになっている。この携帯端末を用いて、105のように物体に記載されたURLアドレスや文字行画像を採取し、文字行記載領域の抽出、文字行認識を行い106のような無線通信回線を用いてデータを転送し、107のようにネットワークに接続された計算機を用いて、ネットワーク上の対象コンテンツを選択し転送、表示する。表示されたコンテンツを操作者が参照し、再度指示を行うことにより、双方向の通信が行われる。図10と図3に本発明の画像表示例を示す。図2は撮像された物体に記載されたURL文字行に対し文字認識処理を行い、その認識結果からURLに対応したホームページを受信し表示した例である。まず、201の画像表示部で撮像素子により採取された物体を表示し、画像表示部内には202のようにマーカーが表示されている。マーカーは操作者が指示する事により移動可能である。このマーカーに最も近い位置に記載されているURL文字行に対し、文字行矩形抽出処理が連続して行われ、203のように抽出結果の矩形が随時更新されて表示が繰り返される。このことにより、操作者は端末を動かすか、マーカーを移動させる事により、認識対象文字行を容易に選択する事が可能となる。目的の文字行を矩形が正しく囲んだ際にユーザーが確定指示を行い、その後、確定されて矩形内の文字行に対し文字認識が行われ、204のように認識結果がポップアップで表示される。この認識結果が妥当な場合、操作者が確定指示を行い、その結果、205のようにネットワーク上の計算機から認識結果のURLに対応したコンテンツが携帯端末に転送される。その後、操作者は表示部のコンテンツを参照し、ボタン操作か206のような表示部のボタンの接触操作により、双方向の通信が可能になる。図3は撮像された物体に記載された文字行に文字認識処理を行い、その認識結果を元にネットワーク上のコンテンツを検索し、検索結果を表示した例である。図2の場合と同様に、301で撮像された画像表示し、302のマーカーに最近傍の文字行が抽出され、303のように連続表示された文字行矩形をユーザーが確定指示すると、認識処理の後に304のように認識結果がポップアップで表示される。この認識結果が妥当な場合、操作者は確定指示を行う。その後、認識結果をネットワーク上の検索エンジンに入力し、得られた検索結果を携帯端末が受信し、305のように表示する。操作者は、表示された検索結果を参照し、ボタン操作か表示部の接触操作により、新たにコンテンツを受信することが可能である。図4に携帯端末内のハードウェア構成を示す。携帯端末内には、制御や画像処理、文字認識を行うCPU(401)、メモリ(402)、表示部(403)、入力データを採取するため撮像素子(404)、無線通信に用いる通信部(405)、スピーカー(406)、マイク(407)、電源部(408)、ボタン(409)、表示部に重ねて組み込まれるタッチパネル(410)がバスで結合されている。図5に本発明にお

5

ける全体処理フローを示す。まず携帯端末において、撮像素子を用いたCCDによって501で対象物体を撮像し、撮像した画像を501において表示部で操作者に表示する。その後、図2、図3で示したようにマーカーから最近傍の文字を核として、502で文字行抽出処理を行う。その結果得られた抽出結果を矩形として503で表示部に画像と重ねて表示し、操作者が504で確定指示を行うまで501の画像採取に戻り、繰り返す。操作者が確定指示を出した後、表示された矩形内の部分画像を用いて画像処理を行い、認識処理に入力する。撮像素子により採取した低解像度の入力画像に対し、505によって高精細化処理を行い、高精細化された部分画像を、506の文字行認識に入力する。その後、文字認識結果を507で操作者に提示し指示を仰ぎ、操作者が妥当な認識結果であると判断した場合、確定指示を行い、無線通信によりネットワーク上の計算機に転送される。妥当な認識結果でないと操作者が判断した場合は、501で撮像素子から再度画像を採取して処理をやり直す。計算機に送られた文字行は、509で受信され、510で文字行がURLアドレスか、その他の文字行か判別される。文字行がURLアドレスであった場合、511においてインターネット上でそのURLアドレスのホームページコンテンツを採取し、512で端末に送信する。次に513で文字行がE-mailアドレス化判別し、E-mailアドレスの場合、514でそのアドレスを宛先フィールドに挿入したメール送信プログラムを端末に送信する。このメール送信プログラムはJ A V A（登録商標）等により記述されたアプリケーションを用いることで実現できる。また一度送信されたメール送信用プログラムは携帯端末内に保管され、再度用いることも可能である。また、文字行がURLアドレスでもE-mailアドレスでもない場合、515でインターネット上のコンテンツ検索エンジンに入力し、文字行をキーワードとして検索を行う。その検索結果を、516で端末に送信し操作者に提示する。携帯端末側では517でコンテンツを受信し、操作者が518でコンテンツを参照し、表示部の接触指示かボタン操作により、希望のコンテンツを選ぶか、メール送信等を行い、目的を達成する。また、上記の携帯端末とネットワーク上の計算機とのデータ転送手順と、処理分担は他にも考えられる。例えば、携帯端末上で対象となる文字行記載領域を選択した後、部分画像のみネットワーク上の計算機に転送し、計算機上で画像処理と文字認識を行う。この転送手段としては、メールに部分画像を添付して送信する手段などが考えられる。文字認識処理の結果得られた複数の候補を元に、候補順位順にネットワーク上のコンテンツを検索する。複数候補を用いて検索し、妥当なコンテンツが見つかった場合、そのコンテンツを端末に送信し表示する。また、全ての候補を用いて検索しても妥当なコンテンツが見つからなかった場合、再度パラメータを変えて文字認識処理を繰り返す。上記処理手順を用いれば、操作者が文字記載領域確定指示のみに

10

よって、希望のコンテンツを参照可能なる。ただし、文字認識結果が誤っていた場合でもコンテンツが存在すると、操作者の意図と異なったコンテンツが表示される場合がある。また、上記のような携帯端末と計算機との送受信の際に、課金を行うサービスも考えられる。本サービスは、ネットワーク上に繋がった計算機を管理する会社が、携帯端末から認識結果である文字列や、認識対象となる部分画像を受信し携帯端末側にコンテンツを送信する際に、操作者に対してサービス料として課金を行う。操作者は、対象となる文字列の認識結果を確認後、コンテンツを受信することが可能になり、安定したサービスが可能になる。図6に文字行抽出処理の概念図を示す。まず、601が携帯端末のカメラで採取した画像であり、602は対象文字行が記載されている物体である。画面の真中の603はマーカーであり、このマーカーの位置から最近傍に記載された文字行を抽出する処理が行われる。マーカーは操作者により、タッチパネルかボタン操作により画面上の位置を移動させる事が可能である。次に、604は入力画像を2値化した後、黒画素の塊である連結成分を生成し、それぞれの連結成分の外接矩形を表示した画像である。処理高速化のために、文字行抽出処理は連結成分の外接矩形位置情報を用いて行う。次に、マーカーの最近傍の位置にある連結成分の外接矩形を選択し、605のように太い矩形で表している。この連結成分を文字行抽出の初期状態の「核」として用いる。この「核」となる連結成分は操作者には表示されない。606は文字行抽出の途中経過を表しており、607で示した「核」を元に横方向に近傍の連結成分を統合していくことにより、文字行矩形を形成する。次に607は横方向に統合した結果、連結成分により構成された文字行矩形を表示している。608は操作者から見た文字行矩形抽出結果の表示であり、点線の矩形が文字行抽出結果である。この矩形を操作者が参照し、対象文字行を矩形が正しく囲んでいる場合、確定指示をボタンで行い、そうでない場合は端末を動かすかマーカーを移動させることにより、随時文字行抽出結果矩形が変動する。図7は、文字行抽出処理の処理フローを示している。文字行抽出処理は、まず撮像素子から701で画像を採取し、入力画像を702で2値化し、次に703でノイズ除去を行う。その後、704で2値画像から連結成分の生成を行い、705でマーカーから最近傍の連結成分を初期値の「核」として選択する。次に706で「核」となる連結成分からある一定の近距離の連結成分を統合して、文字行矩形を形成し、矩形座標が outputされる。この矩形を707の表示部において動的に表示し、708でユーザーが確定指示を出すまで701の画像採取へ戻り処理を繰り返す。ユーザーが確定指示を出した場合、高精細行抽出処理へ矩形座標が入力される。図8は、高精細行画像生成の概念図を示している。文字行画像を認識処理に入力する前に、対象となる部分画像に対し画素補完を行い、高解像度の部分画像を生成

20

30

40

50

6

し、認識精度の向上を狙う。801は操作者が文字行矩形を確定指示した時点での表示部を示している。この文字行矩形位置情報を元に、画像採取時に得られた濃淡画像から部分画像をくり抜いたのが802の画像である。この濃淡部分画像に対し画素補完処理を行い、高解像度化したのが803の画像である。このように画素補完により高解像度化を行うと、斜め線分に階段状の形状が現れる。そこで、画像に平滑化を施すと、804のようにノイズが減り、斜め線分の輪郭も滑らかになる。次に、2値化処理を施すと805の画像のようにノイズも無く、輪郭も鮮明な高精細文字パターン画像が生成される。この高精細行画像を認識処理に入力することにより、低解像度の撮像系を用いても高精度な認識結果が得られる。図9に高精細行画像生成の処理フローを示す。まず文字行抽出処理で決定された文字行矩形の座標を元に、901で画像採取時に得られた濃淡画像から部分画像をくり抜く。部分画像に対して、902で画素補完により画像を拡大する。次に、903で拡大した画像に対し平滑化を行い、904で2値化を施す。その結果、高精細な2値画像が生成される。図10に文字行認識処理の概念図を示す。1001は入力された文字行画像であり、この文字行画像に対し連結成分ごとに文字パターン候補の組み合わせを生成しネットワーク表現したのが1002である。1003は「h」のみ切出したパターンで、1004は「ht」を組み合わせたパターンである。これらの組み合わせのうち、正しい文字パターンを選んでネットワーク上をルート探索すると、文字行の認識結果が得られる。このネットワーク上のパターンを文字認識モジュールに入力した結果が、1005である。1003の「h」のパターンには1006のように「h」の認識結果が得られ、1004の「ht」の2つ文字を組み合わせたパターンに対する認識結果は1007のように「n」になっていることが分かる。このように全てのパターンに対する文字認識結果と、文字認識用いたテンプレートとの類似度をネットワーク上に記録しておく。1008は、認識結果の記録が付加されているネットワークに対し、単語知識と認識結果の類似度を元にルート探索をした結果である。この例ではURLとして頻発する単語（http://, www, co.jp, comなど）は単語知識でルートの制約を行い、それ以外の単語は認識結果の類似度を指標としてルートを決定する。単語知識はDNS（ドメインネームサーバ）からURLアドレスを得て活用するか、看板等を対象物とする際には地理的位置情報を用いて存在単語を限定する手段も考えられる。決定されたルート上の文字をつ

なぎ合わせたのが1009の文字行であり、これが文字行画像に対する認識結果である。図11に文字行認識の処理フローを示す。まず、入力された文字行画像に対し、1101で文字パターンの候補になる組み合わせを生成しネットワークを構築する。次に1102でそれぞれの文字パターンの候補に対し文字認識を行い、それぞれに対応した認識結果と類似度を情報として付加しておく。次に1103で予め蓄積されている単語知識を用いて、ネットワーク上のルート制約を行う。次に1104で単語知識制約と、文字認識結果の類似度を用いてネットワーク上の最適なルートを決定する。その結果、文字行の認識結果が得られる。上記の手段を用いることにより、操作者が着目する文字行を選択し、文字認識結果から情報コンテンツを得ることが可能になる。

【0006】

【発明の効果】本発明を用いると撮像素子を内蔵した携帯端末を用いて、任意の文字行が記載された物体を撮影し、その文字行を元にした情報を入手する事が可能になる。また、任意の文字行を操作者が選択する際に、携帯端末の操作を容易にする事が出来る。また、解像度の低い撮像素子を用いた場合でも、擬似的に高解像度画像を生成し文字認識処理を行う事により、認識精度を高める事が可能となる。

【図面の簡単な説明】

【図1】認識機能付き携帯端末を用いた情報収集の概念図である。

【図2】本発明を用いた携帯端末の表示画面の図である。

【図3】本発明を用いた携帯端末の表示画面の図である。

【図4】携帯端末のハードウェア構成図である。

【図5】携帯端末内とネットワーク上計算機の処理フローの図である。

【図6】文字行抽出の概念図である。

【図7】文字行抽出の処理フローである。

【図8】高解像度画像生成の概念図である。

【図9】高解像度画像生成の処理フローである。

【図10】文字行認識の概念図である。

【図11】文字行認識の処理フローである。

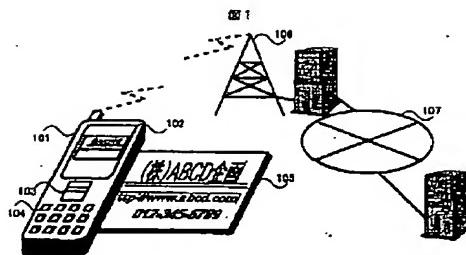
【符号の説明】

201・・・画面表示部 202・・・マーカー 203・・・

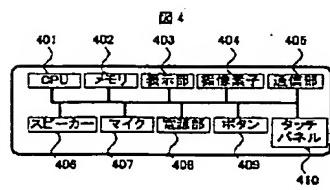
文字行抽出矩形 204・・・文字行認識結果 205・・・

画面表示部 206・・・操作キー。

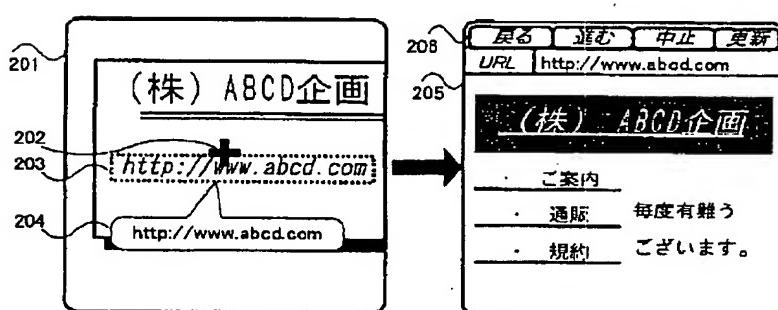
【図1】



【図4】



【図2】



【図1】

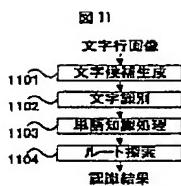
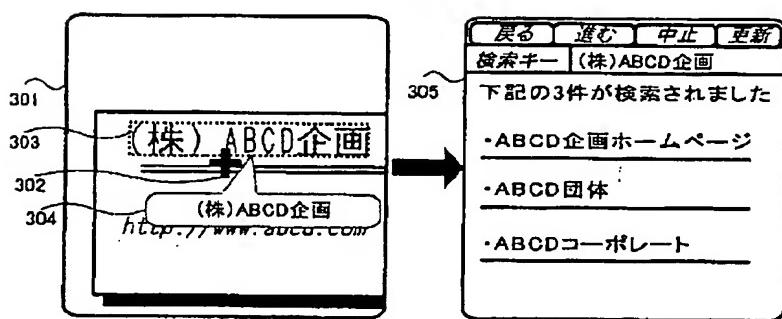


図2

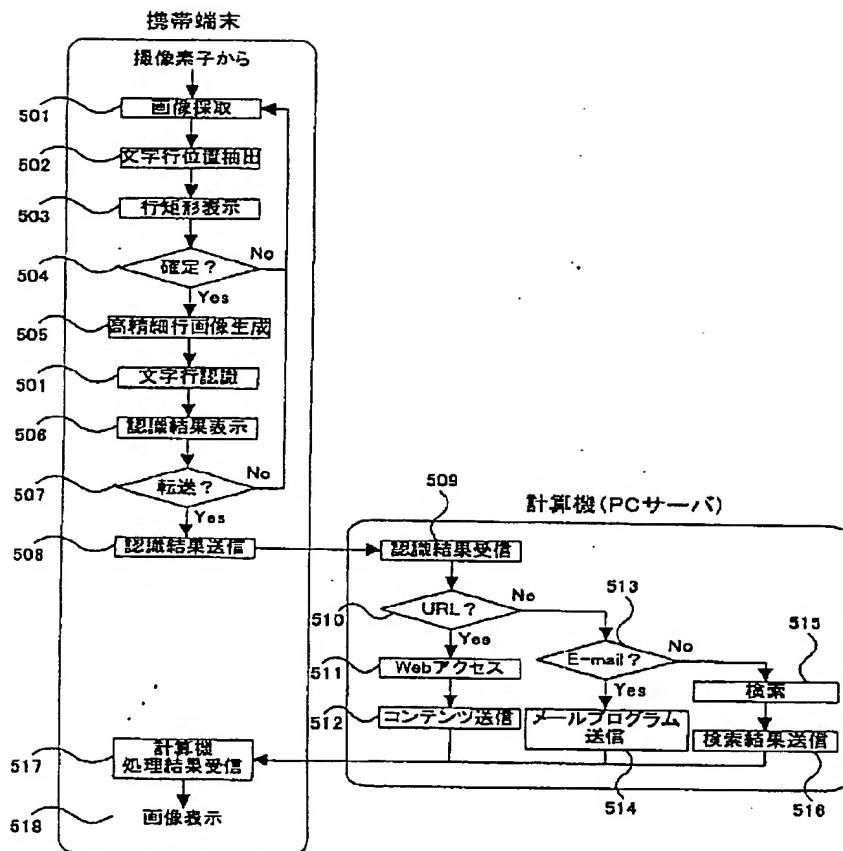
【図3】

図3

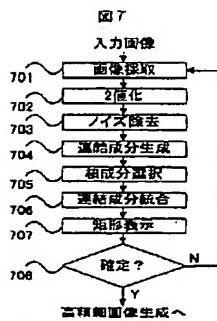


【図5】

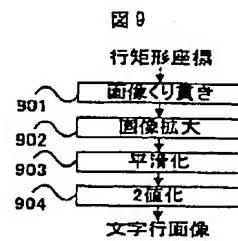
図5



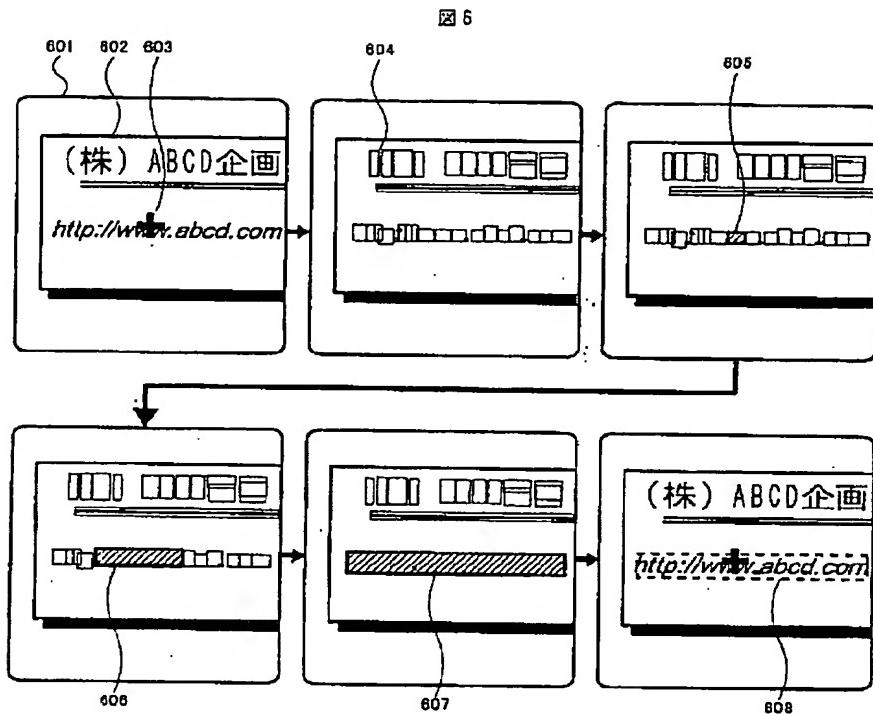
【図7】



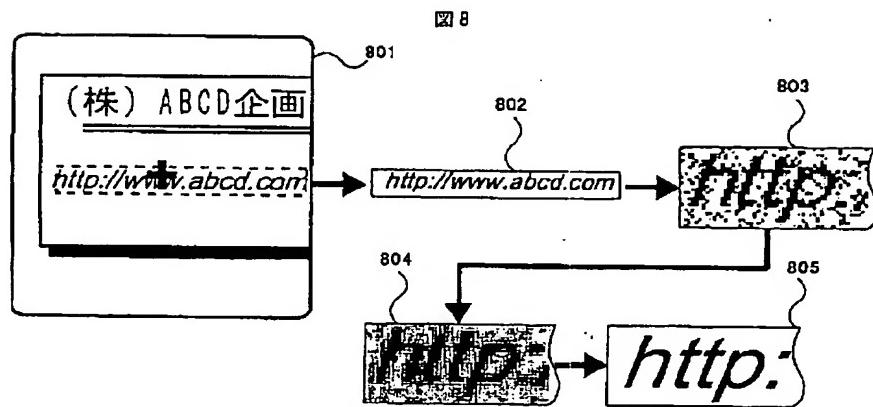
【図9】



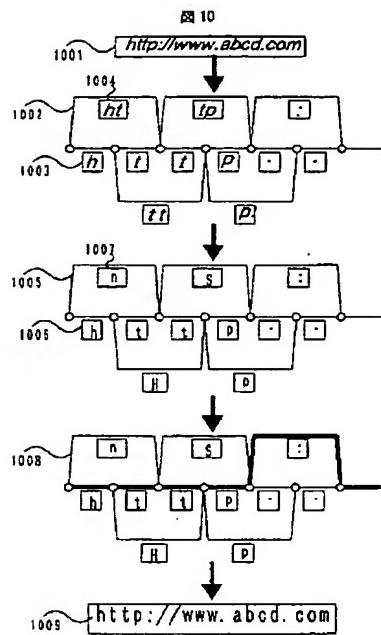
【図6】



【図8】



【図10】



フロントページの続き

(72)発明者 酒匂 裕
 東京都国分寺市東恋ヶ窪一丁目280番地
 株式会社日立製作所中央研究所内

F ターム(参考) 5B064 AA07
 5K027 AA11
 5K067 AA21 BB04 BB21 DD52 DD53
 EE02 EE10 FF02 FF23 FF32
 HH22 KK15
 5K101 KK02 LL12 NN06